

Whitepaper

DATA-CLEANROOMS

WAT ZIJN HET EN HOE ZET JE ZE EFFECTIEF IN?

Januari 2024



INLEIDING

Het landschap van digitale marketing is flink veranderd. Firefox en Safari ondersteunen geen cookies van derden meer, Chrome is ze aan het uitfaseren en de ePrivacy Richtlijn en GDPR vereisen dat uitgevers en adverteerders toestemming van consumenten krijgen om hun data te verzamelen en te gebruiken. En dat zorgt voor de nodige uitdagingen op het gebied van data en privacy. Gelukkig bieden data-cleanrooms een oplossing.

Cleanrooms maken het voor bedrijven mogelijk om samen te werken en data te analyseren met behoud van privacy en naleving van regelgeving. Ze zijn vooral nuttig in situaties waarin meerdere partijen de collectieve kracht van hun gegevens willen benutten zonder de individuele privacy of gegevensbeveiliging in gevaar te brengen. Klantgegevens worden zorgvuldig geanonimiseerd en ontdaan van alle persoonlijke of identificeerbare informatie. Deze gegevens kunnen vervolgens worden geanalyseerd en gebruikt om nieuwe inzichten te krijgen over gebruikers op een geaggregeerd niveau.

Dit whitepaper geeft marketeers die nog geen of weinig ervaring hebben met data-cleanrooms inzicht in alle mogelijkheden. Van het definiëren van doelstellingen tot juridische zaken, uploaden van data en verbinding met andere bronnen tot het uitvoeren van data-analyses. De theorie wordt aangevuld met cases uit de Nederlandse reclame-industrie en aanbevelingen die betrekking hebben op de beperkingen en kosten van data-cleanrooms.

Bijdragen van: Birgit Boing, Emmelijnn Hermens, Weiwei Liu-Schröder, Bastiaan Spaans, Bernard de Vreede, Mark Wilmont



1. WAT IS EEN DATA-CLEANROOM?

In een wereld waarin first-party data in de schijnwerpers staan maar geanonimiseerd moeten blijven, krijgt het concept van data-cleanrooms logischerwijs veel aandacht. Ze zijn trending geworden, omdat ze voldoen aan de behoeften van marketeers: het begrijpen en inzetten van je eigen klantgegevens zonder deze te hoeven delen en zonder het schenden van wetten. Het niet naleven van regelgeving, ongeacht of deze globaal of lokaal is, kan namelijk leiden tot boetes, maar ook tot negatieve publiciteit.

Data-cleanrooms klinken veelbelovend in die zin dat ze kunnen helpen bij het combineren van gegevens uit verschillende bronnen om doelgroepen te begrijpen, terwijl ze ontworpen zijn met het oog op privacy. Dit betekent dat ze garanderen dat alle persoonlijk identificeerbare informatie (PII) wordt opgeslagen en gecontroleerd op privacyregels, door de eigenaar van de gegevens. Dit ook om de op pixel gebaseerde tracking en meting van derden te vervangen. Dit proces maakt het mogelijk om op een veilige manier inzichten te verkrijgen op basis van de gecombineerde gegevens, zonder de privacy van individuele gebruikers in gevaar te brengen.

Met name in de Europese Unie zijn data-cleanrooms een veelbesproken onderwerp vanwege de strenge wetgeving. Desondanks is de populariteit ervan tot nu toe vrij laag. Waarom is dat zo? Ten eerste is er nog steeds niet veel kennis over de werking en de potentiële bescherming die ze bieden. Ten tweede zijn de kosten vaak onduidelijk en ten slotte zijn er nog geen standaarden ontwikkeld.

Toegevoegde waarde van data-cleanrooms

Voor marketeers zijn data-cleanrooms vooral een vertrouwde omgeving waar hun first-party data wordt geüpload, bewaard, gematcht, geanalyseerd (voor bijvoorbeeld overlap/inzichten) of gedeeld voor targetingdoeleinden. Dit alles wordt door een onafhankelijke partij gedaan. Deze partij fungeert als brug tussen de bestanden en betrokken partijen, bijvoorbeeld wanneer een adverteerder en een publisher, of een data-owner, een adverteerder en een publisher, hun data met elkaar willen matchen.

De data wordt door de onafhankelijke beheerder van de data-cleanroom zorgvuldig geanonimiseerd en ontdaan van alle persoonlijke of identificeerbare informatie. Vervolgens wordt deze data geraadpleegd en geanalyseerd, zodat de betrokken partijen op geaggregeerd niveau vernieuwende inzichten kunnen krijgen.

2. DE VERSCHILLENDE FASES VAN EEN DATA-CLEANROOM

Data in een data-cleanroom doorloopt altijd een aantal fases:

1. Anonimisering

Het proces in de data-cleanroom start met anonimisering en aggregatie van de datasets. Hierbij wordt alle persoonlijk identificeerbare informatie (PII) uit de oorspronkelijke datasets verwijderd of gecodeerd, waardoor individuele privacy wordt gewaarborgd.

2. Data-aggregatie

In de tweede fase volgt data-aggregatie. Dit houdt in dat gegevens uit diverse bronnen worden samengevoegd tot uitgebreidere datasets met een breder bereik en verbeterde inzichten.

3. Toegang tot en analyse van gegevens

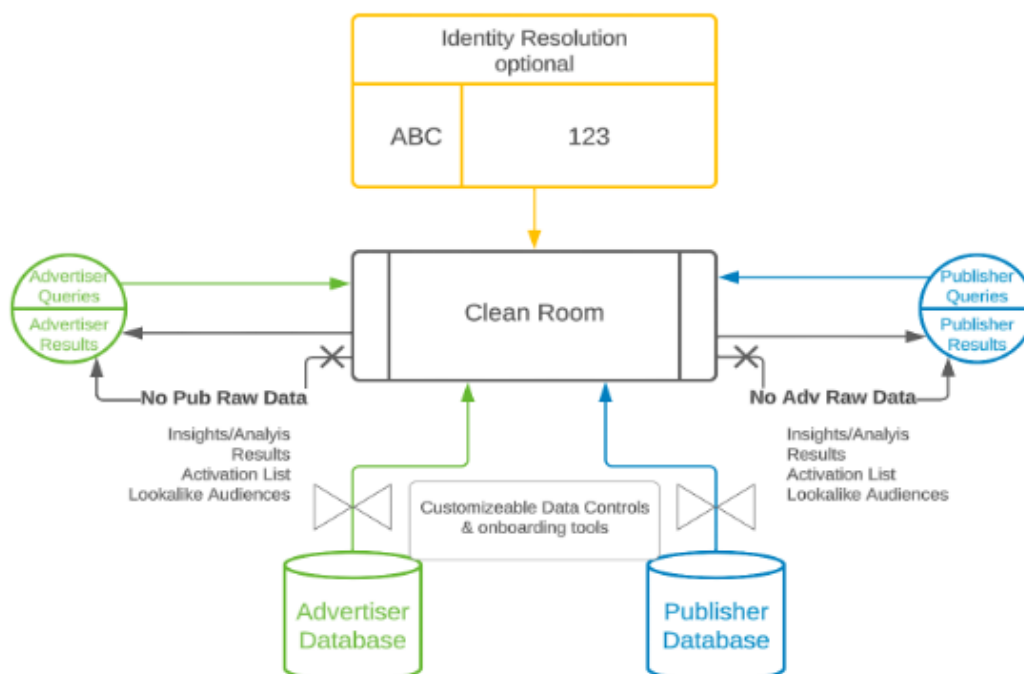
Nadat de data is geanonimiseerd en geaggregeerd, kunnen gebruikers beginnen met het uitvoeren van query's en analyses.

4. Inzichten en samenwerking

Binnen de cleanroom kunnen deelnemers samenwerken, inzichten delen en bevindingen uitwisselen. Deze samenwerking stelt iedereen in staat hun kennis en expertise te bundelen. Uitgevers kunnen bijvoorbeeld inzicht bieden in gebruikersgedrag, terwijl adverteerders inzicht verschaffen in de prestaties van campagnes. Deze gezamenlijke inspanning maakt het mogelijk targetingstrategieën te verfijnen, campagnes te optimaliseren of nieuwe advertentiebenaderingen te ontwikkelen.

5. Activatie/destinations

In de laatste fase kunnen de meest relevante doelgroepsegmenten eenvoudig worden geïdentificeerd en vervolgens direct 'geactiveerd' (targetingdoeleinden) bij uitgevers en walled gardens of via transactieplatformen zoals DSP's, SSP's, en ad servers. Als de data de bestemming bereikt heeft, is het aan de uitgever of adverteerder zelf om de audience aan de campagne toe te voegen voor activatie of het uitsluiten van eigen klanten.



Schema gemaakt door Travis Gossen (Platforms Sales Engineer, Google)

3. HOE BEGIN JE?

Nu duidelijk is wat een data-cleanroom precies is en welke fases hij doorloopt, is het moment aangebroken om zelf ook aan de slag te gaan met data-cleanrooms. Maar waar te beginnen? Een handig stappenplan:

1. Marktonderzoek en analyse

- Onderzoek welke PII-data intern aanwezig is.
- Neem contact op met zowel de IT- als de juridische afdeling en vraag of er consent van de gebruikers/klanten is om de data te gebruiken voor marketing- en advertentiedoeleinden?
- Maak een use-case en bedenk wat je daarmee wil bereiken. Vergeet hierin niet match-rates en potentieel bereik als KPI's mee te nemen voor de schaalbaarheid van je campagnes.
- Maak een overzicht welke data-cleanroomaanbieders er in Nederland actief zijn en bekijk of hier aanbieders met lokale proposities tussen zitten. Denk aan de volgende voorwaarden/criteria/aspecten:
 - Heeft de data-cleanroom een integratie/samenwerking met jullie preferred marketingplatformen (DSP/SSP/DMP/CPD)?
 - Zijn er lokale uitgevers of andere lokale data-aanbieders al aangesloten als partners?
 - Heeft de data-cleanroom ook activatiemogelijkheden met de walled gardens van Google en Meta?
- Stel een team samen van interne stakeholders. Het is altijd verstandig om hierin meerdere disciplines of afdelingen op te nemen. Wijs bijvoorbeeld een product-/data-eigenaar van het marketing- en/of CRM-team aan en laat die nauw samenwerken met de data-technologen die het gebruik overzien.

2. Doelstellingen bepalen

- Bepaal het specifieke doel van het delen en analyseren van de gegevens in de cleanroom.
- Stel de gewenste inzichten of analyses in.

3. Juridische overwegingen en toestemming

- Zorg ervoor dat alle betrokken partijen juridisch in overeenstemming zijn over het delen van de gegevens.
- Kom een data-processor-agreement overeen waarin duidelijk staat omschreven wat het doel en de middelen zijn. Kortom, welke data gaat er verwerkt worden, bijvoorbeeld e-mailadres, huisadres en telefoonnummer. Maar ook online domeinen zoals IP, ID of user agent.
- Verkrijg eventuele vereiste toestemmingen of goedkeuringen voor gegevensdeling.

4. Zelf doen of uitbesteden?

- Ga je zelf je eigen data-cleanroominfrastructuur bouwen of kies je ervoor om samen te werken met een externe partij? Beide opties hebben voor- en nadelen en variëren in kosten. Doorgaans is het bouwen van een eigen infrastructuur kostbaar en vooral gunstig voor zeer grote organisaties met omvangrijke datasets.

De kosten van het gebruik van data-cleanrooms variëren, afhankelijk van de datasets. Als richtlijn kun je 15% van je totale campagnebudget toewijzen aan audience matching en ongeveer 10% van het totale campagnebudget reserveren voor infrastructuur.

5. Prijsstelling

- Kies je voor uitbesteden? Stel dan wel een licentieovereenkomst met de aanbieder van de data-cleanroom op. Bij data-cleanrooms in het marketingdomein gaat dat meestal op basis van het aantal records under management (aantal CRM-records) en de hoeveelheid collaboratiepartners en activatie-channels waarmee je wil gaan samenwerken om de juiste audiences te activeren.

6. Data uploaden

- Upload de geanonimiseerde gegevens naar de beveiligde omgeving van de data-cleanroom.
- Zorg voor versleuteling en veilige overdracht van gegevens om de integriteit te waarborgen.

7. Verbinden met andere bronnen

- Integreer, indien nodig, de geüploade gegevens met andere relevante bronnen.
- Creëer een geïntegreerde dataset voor een holistisch beeld.

8. Toegangsbeheer

- Implementeer strikte toegangscontroles om ervoor te zorgen dat alleen geautoriseerde gebruikers de gegevens kunnen bekijken en analyseren.
- Monitor en registreer toegangsacties voor veiligheidsdoeleinden.

9. Data-analyse

- Voer de gewenste analyses uit binnen de gecontroleerde omgeving van de cleanroom.
- Analyseer gegevens op geaggregeerd niveau zonder individuele privacy te schenden.

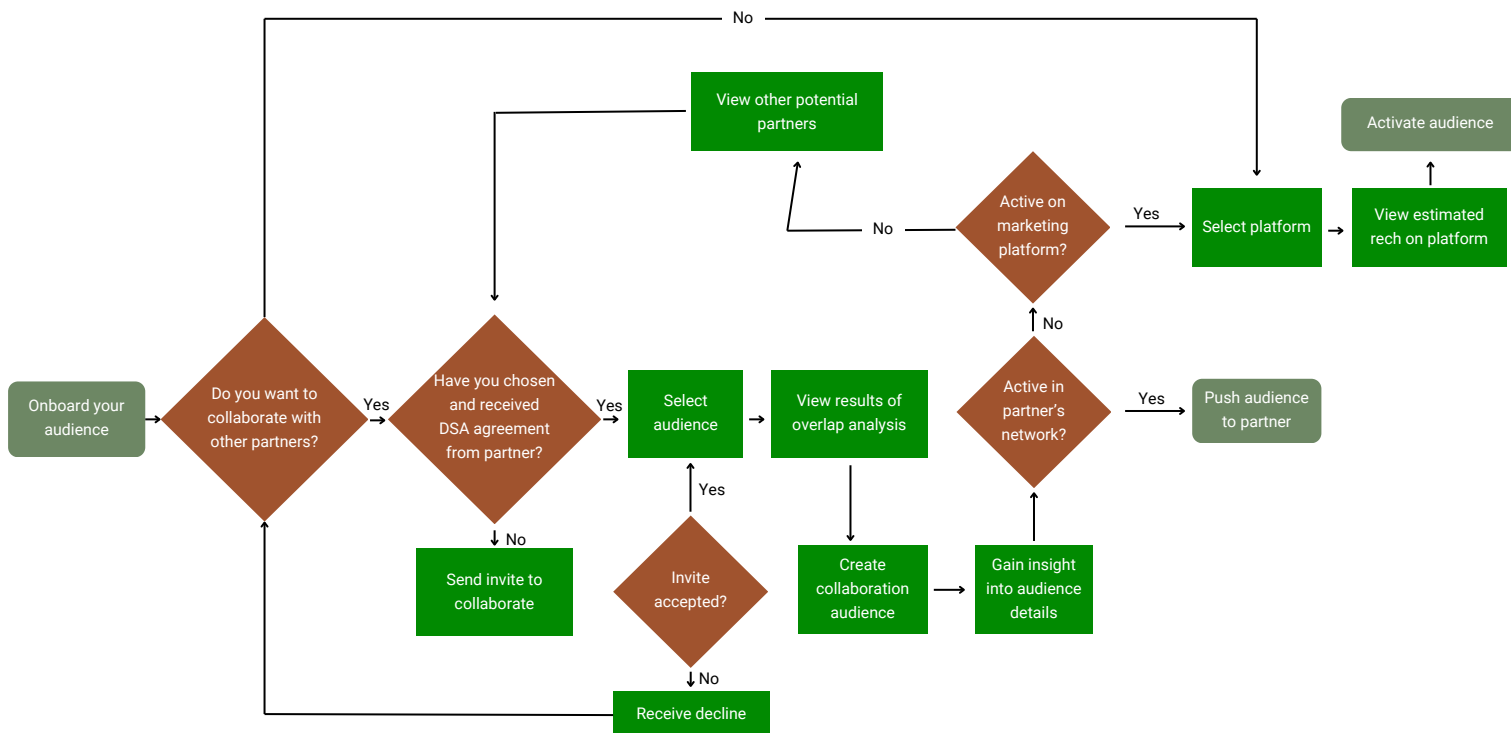
10. Resultaten delen

- Deel de verkregen inzichten op een geaggregeerde en privacyveilige manier met de betrokken partijen.
- Zorg ervoor dat het delen van de resultaten voldoet aan regelgeving en privacystandaarden.

11. Beveiliging en monitoring

- Handhaaf voortdurend de beveiligingsmaatregelen om de integriteit van de gegevens en de veiligheid van de cleanroom te waarborgen.
- Monitor activiteiten om verdachte of ongeautoriseerde toegang te identificeren.

Als je bovenstaande hebt ingeregeld, kun je gaan onboarden. Op de volgende pagina vind je een schematische weergave van hoe dat in zijn werk gaat.



Technisch stappenplan onboarding-proces is gemaakt door Svitlana Fitsak (Product Owner, digitalAudience)

4. PRAKTIJKVOORBEELDEN

Er zijn honderden cases waarin data-cleanrooms een belangrijke rol hebben gespeeld. Uit al deze cases hebben wij een selectie gemaakt van vier merken die allemaal hun eigen beweegredenen, doelgroep en doelstelling hadden om een cleanroom in te zetten. In plaats van de cases volledig uit te schrijven, verwijzen we met behulp van links naar externe pagina's waar de cases op staan en verschillende praktische implementaties getoond worden.

VOORBEELD 1

Booking.com heeft Googles Ads Data Hub (ADH) gebruikt om inzicht te krijgen in de prestaties van hun campagnes en tegelijkertijd de privacy van hun klanten te beschermen. Het bedrijf was in staat om belangrijke informatie te verzamelen, verkrijgen en genereren die kan worden gebruikt om betere beslissingen te nemen over het stimuleren van incrementele gebruikersacquisitie.

[BEKIJK CASE](#)

VOORBEELD 2

In deze case lees je hoe DELA op advertentiekosten heeft bespaard door met behulp van een data-cleanroom bestaande klanten te leren herkennen in een nieuw Customer Match-proces

[BEKIJK CASE](#)

[VIDEO](#)

VOORBEELD 3

Delta Fiber verhoogde hun match-rate met maximaal 68% op Google, Meta en bij uitgevers met digitalAudience.

BEKIJK CASE

VOORBEELD 4

DPG en TUI België hebben data-cleanrooms ingezet met als effect dat ze een meer geëngageerd publiek aantrokken.

BEKIJK CASE

5. AANBEVELINGEN, DISCUSSIE EN TOEKOMSTBEELD

Zoals aan alles in het leven, zitten er ook aan het gebruik van data-cleanrooms voor- en nadelen. We hebben de belangrijkste op een rij gezet:

Voordelen

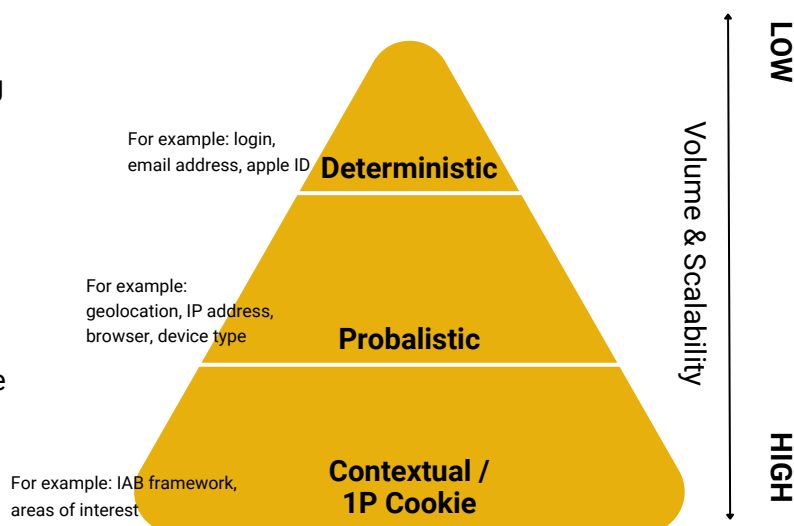
- Audience matching biedt inzichten die niet kunnen worden verkregen uit interne bedrijfsdata.
- Matching gebeurt op een privacyvriendelijke manier met expliciete toestemming van gebruikers.
- Datasets kunnen herhaaldelijk worden hergebruikt voor meerdere campagnes.

Nadelen

- Het is een kostbare onderneming, waarbij ongeveer 25% van het campagnebudget wordt besteed aan audience matching, naast de kosten voor producteigenaren en -ontwikkeling.
- Het vereist aanzienlijke schaalgrootheid, met grote datasets van bijvoorbeeld e-mailadressen.
- Er is voldoende verkeer op de website nodig om gepersonaliseerde ervaringen te bieden.

Datapiramide

Naast deze voor- en nadelen zijn er nog tal van andere facetten die je bewust onder de loep moet nemen voor je de keuze maakt om wel of niet met een data-cleanroom te starten. Om je te helpen bij het maken van deze keuze hebben we als taskforce een 'datapiramide' gebouwd. Deze piramide geeft de afweging tussen schaal en volume per datasoort weer.



- **Activatie versus resultaat**

Lever het gebruik van gematchte doelgroepen significant betere resultaten op? En wegen deze resultaten op tegen de kosten?

- **Nauwkeurigheid versus schaal**

Wat wordt beschouwd als een acceptabele match-rate voor lokale audience matching (dus geen Google/Meta)? Hoe groot moet de doelgroep zijn om effectieve resultaten te behalen? Hoe beïnvloedt schaal de nauwkeurigheid van de doelgroep?

- **Niche versus generiek**

Is het altijd nodig om gematchte doelgroepen te gebruiken om specifieke doelgroepen te bereiken, of is aanwezigheid in het algemene ecosysteem voldoende?

Wat kan nog beter?

De voordelen spreken voor zich: data-cleanrooms bieden prachtige oplossingen voor organisaties die met derden willen samenwerken om data te analyseren en in te zetten voor marketingdoeleinden, en dat alles met behoud van privacy en naleving van regelgeving. Maar zoals aan de lijst van nadelen te zien is, er is nog genoeg ruimte voor verbetering. De markt zal dan ook op meerdere vlakken moeten investeren in de doorontwikkeling van data-cleanrooms, zoals:

- Het creëren van standaardprocedures voor audience matching en data-cleanrooms, lokaal of op industrieel niveau, om eerlijke vergelijkingen tussen adverteerders, leveranciers van data-cleanrooms en uitgevers mogelijk te maken.
- Het toepassen van de FAIR-dataprincipes om een nationale infrastructuur te bevorderen in plaats van geïsoleerde eilanden van data.
- Het gebruik van een uniform algoritme in alle data-cleanrooms om consistente matchingmethoden mogelijk te maken.
- Eén uniforme identificatiemethode voor alle Nederlandse media, in een ideale wereld zouden we daar naar streven. Dat zou de sector in staat stellen om op een transparante en efficiënte manier met data om te gaan. Om deze reden is het allicht de moeite waard initiatieven te overwegen zoals een 'NL-profiel 3.0' om de identificatie van Nederlandse media te stroomlijnen.
- Het bevorderen van transparantie in het gebruik van data voor marketingdoeleinden ten behoeve van consumenten, wat zowel het vertrouwen van de consument als de merkvoorkeur ten goede komt.¹ Dat is een wenselijke situatie voor zowel de consument als merken.

Tot slot

We kunnen als taskforce niet anders dan concluderen dat data-cleanrooms interessante mogelijkheden bieden voor marketeers en uitgevers. Het geeft hen de mogelijkheid om kennis en data te bundelen, met diepere, nauwkeurigere inzichten als gevolg. De gedeelde kennis opent niet alleen deuren voor optimalisatie, maar ook voor het ontwikkelen van vernieuwende strategieën die de gehele sector naar een hoger niveau tillen.

Terwijl organisaties zich blijven inzetten voor privacy en wettelijke naleving, fungeert de data-cleanroom als een essentiële schakel in het behouden van een delicaat evenwicht tussen waardevolle inzichten en individuele privacybescherming. Vooruitkijkend belooft deze aanpak niet alleen te voldoen aan de huidige eisen, maar ook de weg vrij te maken voor een toekomst waarin gegevensdeling en samenwerking veiliger en effectiever kunnen plaatsvinden dan ooit tevoren.

¹V. Bruns and M. Minckler, People take their privacy more seriously than you'd expect – brands should too, Google, september 2022